dbGaP Study Release Notes

Release Notes for NHLBI TOPMed - NHGRI CCDG CATHGEN, phs001600.v2.p2
"*NHLBI TOPMed - NHGRI CCDG: CATHeterization GENetics (CATHGEN) Cohort*"

For any questions or comments, please contact: dbgap-help@ncbi.nlm.nih.gov.

March   9, 2021   Version 1 Data set release date
August   20, 2021   Version 2 Data set release date

2021-08-20
Version 2 Data set release for NHLBI TOPMed - NHGRI CCDG CATHGEN now available

This release includes updated phenotype tables, the addition of Freeze 9 whole genome sequences (WGS), and corresponding VCFs. Please refer to the latest study configuration report for a detailed description of each download component.

Consent group 1 (c1): Disease-Specific (Cardiovascular Disease, IRB) (DS-CVD-IRB)

Data Type	subjects	samples
Phenotype	1271	1411
Seq_DNA_SNP_CNV (VCFs)	1265	1405
WGS	1265	1405

For a description of SAMPLE_USE terms, please see:
https://www.ncbi.nlm.nih.gov/projects/gap/submission/GetSampleUseTypes.cgi

Study and Phenotype Updates

1. **New Study Accession**

   NHLBI TOPMed – NHGRI CCDG CATHGEN version 1 phs001600.v1.p1 has been updated to version 2. The dbGaP accession for the current data is **phs001600.v2.p2**. The participant number (p#) has changed in version 2; subjects have been retired. No new subjects have been added to this study.

2. **Updated Datasets (n=4 datasets; all existing variables were updated)**

pht	version	Dataset Name
11019	2	TOPMed_CCDG_CATHGEN_Subject
11020	2	TOPMed_CCDG_CATHGEN_Sample
11021	2	TOPMed_CCDG_CATHGEN_Subject_Phenotypes
11022	2	TOPMed_CCDG_CATHGEN_Sample_Attributes

3. **New Variables (n=1 variable)**

pht	pht version	Dataset Name	phv	Variable Name
11019	2	TOPMed_CCDG_CATHGEN_Subject	492409	SEX

Molecular Data Updates

dbGaP Study Release Notes

1. See download components phg001522.v1.TOPMed_CCDG_CATHGEN.sample-info.MULTI.tar for manifest of genotyped samples and files.
2. Genotype calls are available in matrix format ('genotype-calls-vcf') as VCF file(s) with samples of same data sharing consent.
3. Quality control data are in phg001522.v1.TOPMed_CCDG_CATHGEN.genotype-qc.MULTI.tar.

phg	freeze	sample_cnt	subject_cnt
phg001522.v1	8	123	123
phg001611.v1	9	1405	1265

Authorized Access (Individual Level Data)

Individual level data are available for download through the dbGaP Authorized Access System upon approval of the Data Access Request (DAR):

- https://dbgap.ncbi.nlm.nih.gov/aa/wga.cgi?login=&page=login

Public FTP site (Summary Level Data Only)

All data tables, data dictionaries, and documents will be housed under one directory for ease of downloading. The data_dict filenames have an added study version number (phs#.v#) and deleted participant set number (p#) from the table accession (pht#.v#). The var_report filenames have an added study version number (phs#.v#). In the var_report files, variables contain version numbers (phv#.v#) and summaries were created for each consent group (c#). These FTP files are available at:

- https://ftp.ncbi.nlm.nih.gov/dbgap/studies/phs001600/phs001600.v2.p2

2021-03-09
Version 1 Data set release for NHLBI TOPMed - NHGRI CCDG CATHGEN now available

This release includes the addition of Freeze 8 whole genome sequences (WGS) brokered through the Sequence Read Archive (SRA), and VCFs derived from WGS. Please refer to the latest study configuration report for a detailed description of each download component.

Consent group 1 (c1): Disease-Specific (Cardiovascular Disease, IRB) (DS-CVD-IRB)

Data Type	subjects	samples
Phenotype	1279	1420
Seq_DNA_SNP_CNV (VCFs)	123	123
WGS*	123	123

*These data are brokered through the Sequence Read Archive (SRA). Please see Authorized Access instructions below.

For a description of non-SRA SAMPLE_USE terms, please see:
https://www.ncbi.nlm.nih.gov/projects/gap/submission/GetSampleUseTypes.cgi

dbGaP Study Release Notes

Molecular Data

1. See download components phg001522.v1.TOPMed_CCDG_CATHGEN.sample-info.MULTI.tar for manifest of genotyped samples and files.
2. Genotype calls are available in matrix format ('genotype-calls-vcf') as VCF file(s) with samples of same data sharing consent.
3. Quality control data are in phg001522.v1.TOPMed_CCDG_CATHGEN.genotype-qc.MULTI.tar.

Authorized Access (Individual Level Data and SRA Data)

Individual level data and Sequence Read Archive (SRA) data are available for download through the dbGaP Authorized Access System upon approval of the Data Access Request (DAR):

- https://dbgap.ncbi.nlm.nih.gov/aa/wga.cgi?login=&page=login

Public FTP site (Summary Level Data Only)

All data tables, data dictionaries, and documents will be housed under one directory for ease of downloading. The data_dict filenames have an added study version number (phs#.v#) and deleted participant set number (p#) from the table accession (pht#.v#). The var_report filenames have an added study version number (phs#.v#). In the var_report files, variables contain version numbers (phv#.v#) and summaries were created for each consent group (c#). These FTP files are available at:

- https://ftp.ncbi.nlm.nih.gov/dbgap/studies/phs001600/phs001600.v1.p1